Detlef Dürr
Dustin Lazarovici

Understanding Quantum Mechanics

The World According to Modern Quantum Foundations

Springer

Detlef Dürr
Mathematisches Institut
Universität München
München, Germany

Dustin Lazarovici
Faculté des lettres Section de philosophie
Université de Lausanne
Lausanne, Switzerland

Translation by the authors, editing by Stephen Lyle.

ISBN 978-3-030-40067-5 ISBN 978-3-030-40068-2 (eBook)
https://doi.org/10.1007/978-3-030-40068-2

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Some Mathematical Foundations of Quantum Mechanics

1

> Philosophy is written in that great book that always lies before our eyes – I mean the universe. But one cannot understand it if one does not first learn the language and know the signs in which it is written. This language is mathematics, and the signs are triangles, circles, and other geometric figures, without which it is impossible for man to understand a single word of it; without these one is just wandering around in a dark labyrinth.
>
> Galileo Galilei, *Il Saggiatore*[1]

Here we recall some mathematical basics of quantum mechanics about which there is no dispute. These fundamentals are equally relevant for all quantum mechanical theories (which are unfortunately—or rather, mistakenly—often referred to as interpretations). Our selection is also determined by our needs in later chapters.

The understanding of a physical theory can take place on different levels. The physical worldview must first be communicable to anyone with an honest interest in the subject, i.e., technical details must not be relevant at this level. This means that one should not hide one's own lack of clarity about the worldview behind statements like: "The theory can only be understood by people who have studied physics for at least 4 years". But on a deeper level, the expert level, theoretical (and experimental) background knowledge is necessary for anyone who wishes to have a solid foundation on the basis of which they can explore the intricacies of the description of nature more deeply. The first chapter of this book is meant to provide some key features of that foundation and should be read exactly in this spirit. From a technical point of view, it may be the most challenging chapter, but every physics student should work through the following results carefully at least once in her life. For some, this may mean skipping the more difficult derivations on a first reading and returning to them at a later point in their studies. In any case, the reader should not feel intimidated by the mathematics, as we will try to provide enough context and explanation alongside the technical details.

[1] G. Galilei, *Il Saggiatore*, Capitolo VI. 1623. [Translation by authors.]

D. Dürr, D. Lazarovici, *Understanding Quantum Mechanics*,
https://doi.org/10.1007/978-3-030-40068-2_1

There are three mathematical pillars upon which quantum mechanics is based, but which have at the same time given rise to the century long debate about the meaning of quantum mechanics. They are succinctly summarised in Remarks 1.1, 1.2, and 1.4. But before presenting these, we begin by briefly discussing a non-mathematical term, whose omission is in fact the true source for much of the debate about quantum mechanics.

1.1 Ontology

There are a cornucopia of books about the meaning of quantum mechanics. In most of them notions such as *mystical*, *incomprehensible*, *quantum logical*, *information*, *collapse*, and *observer* pop up almost continually. One particular term, however, hardly ever occurs: *ontology*. The ontology[2] of a physical theory specifies what the theory is about. Since in so-called classical physics it is clear at the outset what the physical theory is about—e.g., Newtonian mechanics is about the motions of point particles—there was no need for an extra Greek word to philosophize about the obvious. But if we wish to understand the confusion about quantum theory, we cannot avoid the term. The reason is simply that, in typical presentations, it is unclear what quantum theory is about. And indeed, each so-called interpretation of quantum mechanics tries to develop its own idea.

John Stewart Bell, who will be mentioned on several occasions throughout this book, invented the term *beables*—a neologism derived from "to be" and "able". Beables are to be contrasted with "observables", or observable quantities. To appreciate the difference, note that observation or measurement is actually a complex physical process. Our measuring devices and sensory organs are complex physical systems that are subject to physical laws and which interact with the measured or observed objects. It is therefore nonsense to think of observed quantities as fundamental in the formulation of a physical theory.[3] With the term *beables*, Bell wanted to stress the fact that a precise physical theory should deal with what there is in the world, i.e., it should deal with the subject of our observations or what causes them. In the mathematical formalism of a physical theory there must therefore be some variables that refer to physical entities out there in the world. These can be particles, fields, strings, or *GRW flashes* (which we shall discuss later)—whatever it is that the theory posits as the elementary building blocks of matter. These elementary objects are the *beables*, as postulated in the ontology of the theory. If the ontology is unclear, then it can never be clear what the theory has to say about the world.

We see a table over there. Why? Because there is a table over there. But physical theories are not about tables as elementary objects. Instead we have an atomistic theory of matter and the table is therefore considered to consist of atoms. We

[2] From ancient Greek, meaning the study of "that which is".

[3] Compare with the quote from Einstein at the beginning of Chap. 8.

can conceive of a theory of atoms in which atoms are the fundamental ontology, or *beables*. The theory then provides a way to understand the physical properties of the table from the behaviour of its constituent atoms (possibly through their interactions with fields): its shape, its weight, its temperature, its solidity, its electrical conductivity, etc., can all be explained in terms of atoms. Of course, it has long been known that what we call atoms are not elementary at all. They themselves consist of smaller building blocks, and in such a "finer" theory, these even more elementary building blocks would form the ontology. Ontology also stands for what we consider as being physically "real" in our world, and it is indeed a painful process to learn that what we take to be real can change as theory progresses.

We shall use the term ontology from time to time. We need it to understand quantum theory, because the quandary of orthodox quantum theory is caused by a simple dogma: quantum theory must not be about ontology. But then what is it about? That is what the cornucopia of books take it upon themselves to discuss. In the chapters to come we shall show that the quandary evaporates once quantum mechanics is based on a clear ontology. In a famous German poem by Christian Morgenstern, Palmström concluded razor-sharply that what must not be cannot be. And in fact there were and still are many attempts to turn "shan't" into "can't". We shall talk about that, too.

1.2 The Wave Function and Born's Statistical Hypothesis

A central element of quantum mechanics is the wave function of an N-particle system in three-dimensional space, i.e., in $\mathbb{R}^3$ (this is the generally accepted way of speaking even in quantum theories in which particles do not occur as entities at all):

$$\psi : \mathbb{R}^{3N} \times \mathbb{R} \to \mathbb{C}\,, \quad \psi(\mathbf{q}_1, \ldots, \mathbf{q}_N, t)\,. \tag{1.1}$$

Here $\mathbb{C}$ is the set of complex numbers, that is, ψ is a complex-valued function which takes as input a time t and N points in $\mathbb{R}^3$, describing a possible configuration of N particles in three-dimensional space. The time evolution of the wave function with potential V obeys the Schrödinger equation, which we write in terms of the *configuration* variable $q = (\mathbf{q}_1, \ldots, \mathbf{q}_N) \in \mathbb{R}^{3N}$:

$$\mathrm{i}\hbar \frac{\partial}{\partial t}\psi(q,t) = -\sum_{n=1}^{N} \frac{\hbar^2}{2m} \Delta_n \psi(q,t) + V(q)\psi(q,t)\,, \tag{1.2}$$

with the Laplace operator $\Delta_n = \partial^2/\partial \mathbf{q}_n^2$.

There is a disagreement about whether wave functions actually exist for systems of very "large" size, e.g., a measuring apparatus in a laboratory, or the laboratory itself, or even the whole universe. This disagreement will be discussed in Chap. 2 on the measurement problem. However, we mention the origin of the disagreement here because it runs through the whole of quantum mechanics.

Remark 1.1 (Superposition Principle) The Schrödinger equation is a *linear* (partial differential) equation. This means that the sum of constant multiples of solutions of the equation is also a solution of the equation. In the usual jargon, we say that solutions can be superposed.

In addition to the Schrödinger equation, there is a second important equation which led Max Born (1882–1970) to the accepted interpretation of the wave function as a probability amplitude, although only after a correction by Schrödinger.[4] In common parlance, Born's statistical interpretation, which is often referred to as Born's statistical hypothesis or Born's rule, can be stated as follows:

Remark 1.2 (Born's Statistical Hypothesis) If a system has wave function ψ, the measured positions of the particles are distributed according to $\rho = \psi^*\psi = |\psi|^2$. Here ψ^* is the complex function conjugate to ψ.

This means that, if $A \subset \mathbb{R}^{3N}$ is a (measurable[5]) subset of the configuration space, then the probability of finding the system configuration Q in A is given by

$$\mathbb{P}^{\psi}(Q \in A) = \int_A |\psi|^2(\mathbf{q}_1, \ldots, \mathbf{q}_N)\, \mathrm{d}^3q_1 \ldots \mathrm{d}^3q_N\,. \tag{1.3}$$

We note that Born's interpretation gives rise to probabilities in quantum mechanics which then appear in quite different forms. The second equation mentioned above, which is central to this statistical interpretation, is usually derived in textbooks by computing $\partial|\psi|^2/\partial t$ using Schrödinger's equation. The reader is encouraged to carry out this derivation using:

1. the product rule for calculating the derivative of a product,
2. the fact that ψ^* solves the complex conjugated form of the Schrödinger equation (1.2), and
3. the fact that the potential V takes real values, so it drops out in the end.

This leads to a continuity equation, the so-called quantum flux equation:

$$\frac{\partial|\psi|^2}{\partial t} = -\nabla \cdot j^{\psi}\,, \tag{1.4}$$

where $\nabla = (\nabla_1, \ldots, \nabla_N)$, $\nabla_k = \partial/\partial\mathbf{q}_k$, and the *quantum flux* $j^{\psi} = (\mathbf{j}_1^{\psi}, \ldots, \mathbf{j}_N^{\psi})$ is given by

$$\mathbf{j}_n^{\psi} = \frac{\hbar}{2\mathrm{i}m}(\psi^*\nabla_n\psi - \psi\nabla_n\psi^*) = \frac{\hbar}{m}\mathrm{Im}\,\psi^*\nabla_n\psi\,. \tag{1.5}$$

[4] Born had first thought of $|\psi|$ as a candidate for a probability density.

[5] In the sense of mathematical measure theory.

Here Im denotes the imaginary part. The usual argument for $\rho = |\psi|^2$ then proceeds as follows. Integrate (1.4) over the entire configuration space $\Gamma = \mathbb{R}^{3N}$, then transform the volume integral into a surface integral on the right-hand side by application of Gauss' theorem, so that the quantum flux gets integrated over a surface $\partial\Gamma$ at infinity, where the flux is zero:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Gamma \rho\,\mathrm{d}^{3N}q = \int_\Gamma \partial_t\rho\,\mathrm{d}^{3N}q = -\int_\Gamma \nabla\cdot j^\psi\,\mathrm{d}^{3N}q = \int_{\partial\Gamma} j^\psi\cdot\mathrm{d}\sigma = 0\,. \tag{1.6}$$

This shows that the integral $|\psi|^2$ over the whole space is preserved in time and $|\psi|^2$ can indeed be taken as a probability density, because clearly, the total probability, the probability of the sure event, cannot change in time. Normalised to unity, it remains forever at unity. Of course, the invariance of the measure is only a necessary condition, not a sufficient condition, to be able to consider $\rho = |\psi|^2$ as a meaningful probability distribution. In textbook quantum mechanics, the Born rule therefore has the status of a postulate whose setting is ultimately only justified by experiment. However, a theoretical justification of Born's rule is possible, and we shall discuss this in Chap. 4.

1.3 The Spreading of the Wave Packet

An important phenomenon associated with the Schrödinger evolution is the spreading of a wave packet. With some basic mathematical knowledge, it is easily explained. The wave function (in the form of a wave packet) of a particle of mass m can (and should) be thought of as a superposition of plane waves, i.e., in mathematical terms, we should consider its Fourier decomposition. A plane wave with wavelength λ and wave number $k = 2\pi/\lambda$ evolves according to

$$\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\hbar k^2 t/2m)}\,,$$

as can be checked immediately using the "free" Schrödinger equation (1.2) for one particle and for potential $V = 0$. Here $\mathbf{k}$ is the wave vector with length $|\mathbf{k}| = k$. The superposition of the plane waves with weights $\hat{\psi}_0(\mathbf{k})$, i.e., the Fourier transform of $\psi(\mathbf{x}, 0)$, yields

$$\psi(\mathbf{x}, t) = \int \hat{\psi}_0(\mathbf{k})\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\hbar k^2 t/2m)}\mathrm{d}^3k\,, \tag{1.7}$$

whence we may assign a group velocity $\mathbf{v}$ to the wave group around a certain value of $\mathbf{k}$:

$$\mathbf{v} = \frac{\hbar\mathbf{k}}{m}\,. \tag{1.8}$$

To obtain this, we differentiate the dispersion relation $\omega := \hbar k^2/2m$ with respect to $\mathbf{k}$ and evaluate at that $\mathbf{k}$-value around which the wave group is centered.

The formula (1.8) was found long before the Schrödinger equation by Louis de Broglie (1892–1987) as a generalisation to "matter waves" of Einstein's T-shirt formula for photons $E = h\nu$. Indeed, $\mathbf{p} = \hbar\mathbf{k}$ is de Broglie's relationship between the wave number and momentum of the particle. The shorter the wavelength, the faster the wave moves. That's one thing to keep in mind. On the other hand, it is known from lectures on analysis, and in particular, the Fourier transform, that the more concentrated a function is in space, the more plane waves with higher k-values occur in the Fourier decomposition of the function. So if we consider the wave function of a particle that is highly localised around a position, then we know very exactly the location of the particle due to Born's interpretation. And the more closely we need to know the particle's position, the more localised the wave function must be. However, the more localised the wave function, the more plane waves with ever higher k-values will be needed to compose that wave function. And then, since the plane waves all have different speeds, the original wave will break up in the course of time into the plane wave parts running at different speeds, at least when the free Schrödinger equation with $V = 0$ governs the motion. The location of the particle at time T will thus be widely scattered, because the individual waves have travelled different distances, and all the more so as the spread in k-values increases.

Every student of physics should carry out the mathematical examination of this spreading effect at least once. It is as fundamental as the derivation of the quantum flux, and that is why we discuss it here. To do so, we first read the integral in (1.7) as an inverse Fourier transform of the product of the functions $\mathrm{e}^{-\mathrm{i}\hbar\mathbf{k}^2 t/2m}$ and $\hat{\psi}_0(\mathbf{k})$. Next we recall from the study of analysis that the product of two functions becomes a *convolution* under Fourier transform:

$$\widehat{f \cdot g}(\mathbf{k}) = \hat{f} * \hat{g}(\mathbf{k}) := \frac{1}{(2\pi)^{3/2}} \int \hat{f}(\mathbf{k} - \mathbf{k}')\hat{g}(\mathbf{k}')\,\mathrm{d}^3k' \,.$$

The same also applies to the inverse transform. Our first mathematical task will thus be to determine this convolution. We compute the Fourier transform of the first function which looks like a Gaussian up to the factor i, noting that the presence of this factor does not change the result that the Fourier transform of a Gaussian is again a Gaussian.

The rigorous calculation is more involved and uses complex analysis. The result is

$$\psi(\mathbf{x}, t) = \int \frac{1}{\left(2\pi\mathrm{i}\frac{\hbar}{m}t\right)^{3/2}} \exp\left[\mathrm{i}\frac{(\mathbf{x}-\mathbf{y})^2}{2\frac{\hbar}{m}t}\right] \psi_0(\mathbf{y})\mathrm{d}^3y \,, \tag{1.9}$$

which is an important representation of the evolution of the wave function for the initial wave function ψ_0. Evaluating the square in the Gaussian, we obtain

$$\begin{aligned}\psi(\mathbf{x},t) &= \frac{1}{\left(\mathrm{i}t\frac{\hbar}{m}\right)^{3/2}}\exp\left(\mathrm{i}\frac{\mathbf{x}^2}{2\frac{\hbar}{m}t}\right)\int\frac{1}{(2\pi)^{3/2}}\exp\left(-\mathrm{i}\frac{\mathbf{x}\cdot\mathbf{y}}{\frac{\hbar}{m}t}\right)\exp\left(\mathrm{i}\frac{\mathbf{y}^2}{2\frac{\hbar}{m}t}\right)\psi_0(\mathbf{y})\,\mathrm{d}^3y \\ &= \frac{1}{\left(\mathrm{i}t\frac{\hbar}{m}\right)^{3/2}}\exp\left(\mathrm{i}\frac{\mathbf{x}^2}{2\frac{\hbar}{m}t}\right)\hat{\psi}_0\left(\frac{\mathbf{x}m}{t\hbar}\right) \\ &\quad+\frac{1}{\left(\mathrm{i}t\frac{\hbar}{m}\right)^{3/2}}\int\frac{1}{(2\pi)^{3/2}}\left[\exp\left(\mathrm{i}\frac{\mathbf{y}^2}{2\frac{\hbar}{m}t}\right)-1\right]\exp\left(-\mathrm{i}\frac{\mathbf{x}\cdot\mathbf{y}}{\frac{\hbar}{m}t}\right)\psi_0(\mathbf{y})\,\mathrm{d}^3y\,.\end{aligned} \tag{1.10}$$

Note that the second summand in (1.10) goes to zero[6] when $t\to\infty$, because

$$\lim_{t\to\infty}\left[\exp\left(\mathrm{i}\frac{\mathbf{y}^2}{2\frac{\hbar}{m}t}\right)-1\right]=0\,.$$

This means that for large times the wave function is given by the first summand in (1.10), viz.,

$$\psi(\mathbf{x},t)\approx\frac{1}{\left(\mathrm{i}t\frac{\hbar}{m}\right)^{3/2}}\exp\left(\mathrm{i}\frac{\mathbf{x}^2}{2\frac{\hbar}{m}t}\right)\hat{\psi}_0\left(\frac{\mathbf{x}m}{t\hbar}\right)\,. \tag{1.11}$$

This can be interpreted as follows. For large times t the wave function will have moved to places $\mathbf{x}$ for which $\mathbf{k}=\mathbf{x}m/t\hbar\in\operatorname{supp}\hat{\psi}_0$, that is, places which are reached by wave groups centered around those wave vector values. Here $\operatorname{supp}\hat{\psi}_0$ is the set of values for which $\hat{\psi}_0\neq 0$, where supp is an abbreviation for "support". Hence, in a sense, $\operatorname{supp}\hat{\psi}_0$ specifies which plane waves are contained in ψ_0, and these then diverge from each other at different speeds according to the de Broglie relation $m\mathbf{v}=\hbar\mathbf{k}$ for the momentum.

We also find that the momentum $\hbar\mathbf{k}$ is distributed according to the probability density $\left|\hat{\psi}_0(\mathbf{k})\right|^2$. This can be seen as follows. Suppose $\psi_0(\mathbf{x})$ is concentrated around $\mathbf{x}=0$. The particles that have reached the position $\mathbf{X}(t)$ at time $t\gg 0$ have therefore moved approximately with the average momentum $\hbar\mathbf{k}=\frac{m}{t}\mathbf{X}(t)$.

[6] Strictly speaking, we should use Lebesgue's theorem of dominated convergence here, because we exchange integration with taking a limit, but rigor does not bring new insights, so let's ignore that here.

Hence, for any (measurable) subset $A \subseteq \mathbb{R}^3$, the momentum distribution reads

$$\begin{aligned}\mathbb{P}^{\psi}\left(\hbar\mathbf{k} \in \hbar A\right) &\approx \mathbb{P}^{\psi}\left(\mathbf{X}(t) \in \frac{t\hbar}{m} A\right) \\ &\approx \left(\frac{m}{\hbar t}\right)^3 \int\limits_{\frac{t\hbar}{m}A} \left|\hat{\psi}_0\left(\frac{\mathbf{x}m}{t\hbar}\right)\right|^2 \mathrm{d}^3x = \int\limits_A |\hat{\psi}_0\left(\mathbf{k}\right)|^2 \mathrm{d}^3k\,, \end{aligned} \tag{1.12}$$

where we have substituted $\mathbf{k} := \mathbf{x}m/t\hbar$ in the last step. Note that everything follows from Born's rule for the position distribution. And in particular note that we consider here an average momentum, i.e., we know that the particle is at time $t = 0$ here and at a much later time there. Then we take the distance between the positions and divide by the time difference. It is this quantity that is measured during a "momentum measurement". The expected value of the momentum $\mathbf{P} = \hbar\mathbf{k}$ is according to its definition

$$\mathbb{E}^{\psi}(\mathbf{P}) = \int_{\mathbb{R}^3} \hbar\mathbf{k} \left|\hat{\psi}_0(\mathbf{k})\right|^2 \mathrm{d}^3k\,. \tag{1.13}$$

This can be rewritten as

$$\mathbb{E}^{\psi}(\mathbf{P}) = \int_{\mathbb{R}^3} \hat{\psi}_0^*(k)\, \hbar\mathbf{k}\, \hat{\psi}_0(\mathbf{k})\, \mathrm{d}^3k\,. \tag{1.14}$$

In lectures on quantum mechanics we are told that the momentum observable in the position representation is given by $\frac{\hbar}{\mathrm{i}}\nabla$, and we can now easily understand what this means from (1.14):

$$\begin{aligned}\hbar\mathbf{k}\,\hat{\psi}_0(\mathbf{k}) &= \hbar\mathbf{k} \int \psi_0(\mathbf{x}) \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}} \mathrm{d}^3x \\ &= \hbar \int \psi_0(\mathbf{x})\, \mathrm{i}\nabla\, \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}} \mathrm{d}^3x \\ &= \hbar \int \big[-\mathrm{i}\nabla\psi_0(\mathbf{x})\big] \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}} \mathrm{d}^3x \qquad \text{(using integration by parts)} \\ &= \widehat{\left(\frac{\hbar}{\mathrm{i}}\nabla\psi_0(\mathbf{x})\right)} \qquad \text{(by Fourier transform)}\,. \end{aligned}$$

By virtue of the Plancherel identity we then get for (1.14)

$$\mathbb{E}^{\psi}(\mathbf{P}) = \int_{\mathbb{R}^3} \psi_0^*(\mathbf{x}) \frac{\hbar}{\mathrm{i}}\nabla\psi_0(\mathbf{x})\, \mathrm{d}^3x\,.$$

This leads to the introduction of the *momentum operator* $\hat{\mathbf{P}} = \frac{\hbar}{i}\nabla$. In "position space" it is a gradient and in "momentum space" a multiplication operator, multiplying the Fourier transform of the wave function by $\hbar\mathbf{k}$.

Remark 1.3 (Perspectives on Heisenberg's Uncertainty Relation) We have seen that the spreading of the wave function is a wave phenomenon which, in combination with Born's statistical hypothesis, leads to an empirical truth which can justifiably be regarded as one of the most significant innovations of quantum mechanics: the more spatially concentrated a wave function, the more quickly it spreads in time. According to Born, the initial high spatial concentration allows a rather exact knowledge of the position, while the spreading leads to a whole range of possible end positions, whose distribution we have associated with the momentum. The relationship between the exactness of position and momentum measurements is described by Heisenberg's uncertainty relation.

Unfortunately, this is often interpreted as meaning that the uncertainty relation implies the impossibility of simultaneous, arbitrarily accurate measurements of momentum and position, but also that it proves that there can be no particles moving on trajectories in quantum mechanics. Such an assertion is of course unwarranted, but as the derivation of the uncertainty relation is usually presented in a highly abstract manner, its true meaning often remains foggy. For example, the mathematically rigorous derivation of the uncertainty relation is based on the commutator of the "position and momentum observables". It is then claimed that the commutator lies at the heart of the uncertainty relation. But before buying that, we need to be clear about the role the "observables" actually play in the theory. We shall clarify this role in Chap. 7.

However, we can already understand from what has been said so far that what is really responsible for the uncertainty relation is the time evolution of Schrödinger's wave function paired with Born's interpretation. And this leads to a different question: how can we justify Born's interpretation? To come to grips with that, we need to understand the so-called quantum theories without observers, such as Bohmian mechanics. In such theories we will explain and justify Born's statistical rule. Finally, we shall understand why the uncertainty relation has nothing to do with the existence or non-existence of particle paths. And the best way to see this is to derive it from a theory in which particles move on trajectories, as we shall do in Sect. 4.3.

1.4 No Mystery: The Double-Slit Experiment

Richard P. Feynman (1918–1988), one of the great physicists of the last century, began his lecture on the double-slit experiment as follows:

> In this chapter we shall tackle immediately the basic element of the mysterious behavior in its most strange form. We choose to examine a phenomenon which is impossible, *absolutely* impossible, to explain in any classical way, and which has in it the heart of quantum

> mechanics. In reality, it contains the *only* mystery. We cannot make the mystery go away by "explaining" how it works. We will just tell you how it works. In telling you how it works we will have told you about the basic peculiarities of all quantum mechanics. (R.P. Feynman, R.B. Leighton and M. Sands. Feynman Lectures on Physics, Volume 3, pp. 17–18, http://www.feynmanlectures.caltech.edu.)

Now we could say that if even Feynman felt that the double-slit experiment was a mystery, then we shouldn't be ashamed if we feel the same way. That's one possible angle. However, another point of view is this: if after almost a century standard quantum mechanics is still unable to explain—in the true sense of "explain"—such a fundamental phenomenon, then there may be something wrong with the theory. The quantum theories we will discuss in this book will clarify the double-slit experiment and other quantum phenomena, in the sense that they not only "describe" the phenomena correctly, but provide a complete picture of how the phenomena come about.

But let's stick to the Feynman quote for now. Feynman says that it is absolutely impossible to explain the double-slit experiment "in a classical way". This is intuitively correct if "classical" refers to the laws of Newtonian mechanics. However, in addition to Heisenberg's uncertainty relation, the double-slit experiment is often quoted as proof that "classical logic" has lost its validity in quantum mechanics. Or that the experiment shows that particle trajectories are simply impossible. Both views are wrong. The possible particle trajectories are shown in Chap. 8, but to come to grips with the discussion at hand the reader is advised to look at Fig. 8.2 and focus for the time being only on the end points of the trajectories. They represent the typical blackening on the screen, which results in the famous "interference pattern".

So what is all this talk about the failure of "classical logic"? Actually, it is rather strange that such concepts as "quantum logic" and "classical logic" ever entered scientific discussions at all, in particular since the double-slit experiment has nothing to do with all this. The double slit experiment can be described as follows. Two (nearby) slits are imaged by a particle beam on a photo screen and the blackening of the screen shows an interference pattern. More precisely, at very low beam intensities, i.e., only one particle is on the way[7] at any given time, or put another way, we send only one-particle waves through the slits, the interference pattern, made up of completely localized, randomly distributed black dots, develops slowly over time.[8] This is perhaps the most important message of all: the interference pattern emerges gradually, wave by wave (or particle by particle), and is nothing more than an accumulation of impact points. Once this has been understood, it opens up a whole catalogue of questions, from which we would like to draw attention to two in particular:

[7] In an ideal situation, we might imagine sending only one particle through the slits each day ...

[8] ... and that could take years.

1. Why does a pointlike spot appear when it is actually a wave that impinges on the screen? The answer closely concerns the measurement problem of quantum mechanics, which we discuss in Chap. 2 .
2. What is the time distribution of the dots? Suppose the one-particle waves are prepared and sent at fixed times, let's say every full hour. Does every dot appear on the screen after the same amount of time? The answer is no. The times are random, with statistics determined by the quantum flux (1.5). This is usually ignored in textbook presentations and there are various ways to justify that. The main reason why this can be ignored from a practical point of view relates to the asymptotic shape of the wave function (1.11), which ensures that the flow lines eventually become straight. When calculating the distribution of dots on the screen, we can in fact replace the real wave pattern behind the slit by a stationary (i.e., temporally unchanging) wave pattern and apply Born's interpretation to this. The exact analysis is mathematically rather involved, however. It can be found in the chapter on scattering theory in *Bohmian Mechanics, The Physics and Mathematics of Quantum Theory*.[9]

We turn now to the issue of logic. First let slit 1 be closed, then the particle can only go through slit 2 and we can make the following statements about this experiment:

$$\text{The particle goes through slit 2 and hits the screen at } \mathbf{x}. \tag{1.15a}$$

$$\text{The corresponding probability is } |\psi_2|^2(\mathbf{x}). \tag{1.15b}$$

Here ψ_2 is the wave function that emerges from slit 2 as a spherical wave. Here we repeat our previous warning: recalling the answer to question (2) above, the true probability for dots appearing on the screen is determined by the quantum flux, but the probability can effectively be computed from the $|\psi_2|^2$ distribution at a fixed (large) time. The same applies to the case when we close slit 2 and open slit 1. In this experiment, we can make the statements:

$$\text{The particle passes through slit 1 and hits the screen at } \mathbf{x}. \tag{1.16a}$$

$$\text{The corresponding probability is } |\psi_1|^2(\mathbf{x}). \tag{1.16b}$$

The experiment in which both slits are open can then be described as follows:

$$\text{The particle passes through either slit 1 or slit 2 and hits the screen at } \mathbf{x}. \tag{1.17a}$$

$$\text{The corresponding probability is } |\psi_1(\mathbf{x}) + \psi_2(\mathbf{x})|^2 = |\psi_1|^2(\mathbf{x}) + |\psi_2|^2(\mathbf{x}) + 2\mathrm{Re}\,\psi_1^*(\mathbf{x})\psi_2(\mathbf{x}), \tag{1.17b}$$

[9] D. Dürr and S. Teufel, *Bohmian Mechanics. The Physics and Mathematics of Quantum Theory*. Springer, 2009.

where Re is the real part and we observe that this is not equal to $|\psi_1|^2(\mathbf{x}) + |\psi_2|^2(\mathbf{x})$. The difference with (1.17b) is due to the interference between the wave function parts ψ_1 and ψ_2 which emerge from slits 1 and 2. This is a typical wave phenomenon. A water wave that passes through a double slit develops the spherical Huygens waves behind the slits, named after Christiaan Huygens (1629–1695), and these then produce the same interference pattern. In the present case the interference term is given by $2\mathrm{Re}\psi_1^*(\mathbf{x})\psi_2(\mathbf{x})$. And this is the bad news, because (1.15a) and (1.16a) are the alternatives in (1.17a), so logically the corresponding probabilities should add up. But they don't!

So we conclude that classical logic fails. Either that or the particle interpretation is nonsense. Or better still, both. It would only be if

$$\mathrm{Re}\,\psi_1^*\psi_2 = 0\,, \tag{1.18}$$

that is, if the interference vanishes, that the probabilities (1.15b) and (1.16b) would sum and ordinary logic would be redeemed.

But the truth is, this is all a lot of fuss over a red herring. All we need to do is to take physics seriously and describe the two situations correctly:

> Slit 1 (2) is closed and the particle goes through slit 2 (1) and hits the screen at $\mathbf{x}$.

These two statements correspond to the physical situations in the first two experiments, but they are *not* the alternatives in (1.17a). This is obvious, because in the latter case both slits are open. This difference in the physical settings should be noted first. Then we may go on and realize that the assertion that the probabilities associated with the first two experiments should "logically" sum up is based on the assumption that the behaviour of the particles passing through slit 1 does not depend on whether slit 2 is open or closed (and vice versa). The first point is that this assumption is not justified, and the second is that it is clearly wrong in quantum mechanics: if both slits are open, then the wave passing through both slits interferes with itself. The resulting waveform is certainly responsible for the shape of the accumulation of the black dots. The only question that remains is: Where do the black *dots* come from? This is indeed a rather disturbing and yet crucial question because, as already stressed, it is intimately connected with the measurement problem, to be dealt with in Chap. 2.

1.5 The Importance of Configuration Space

After linearity (Remark 1.1) and Born's statistical law (Remark 1.2), we come now to the third pillar of quantum mechanics. For this, we go back to the beginning, namely to the wave function of an N-particle system [see (1.1)]. No matter how abstract or mind-boggling it may sometimes seem to be, the wave function $\psi(\mathbf{q}_1, \ldots, \mathbf{q}_N)$ is first and foremost simply a function on the configuration space of

the N particles. Those who believe that the concept of "particle" is physically ill-founded can take this sentence and the word "particle" as a "manner of speaking". A configuration is the collection of N position variables $\mathbf{q}_i \in \mathbb{R}^3$, i.e., $q = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N)$ and the set of all possible configurations is called the configuration space, denoted $\mathbb{R}^{3N}$.

Remark 1.4 (Configuration Space and Entanglement) The wave function of an N-particle system is defined on the configuration space $\mathbb{R}^{3N}$. This means that all N particles share a common wave function. Only in the special case when the wave function is a product $\psi(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N) = \varphi_1(\mathbf{q}_1)\varphi_2(\mathbf{q}_2)\cdots\varphi_N(\mathbf{q}_N)$ can we say that the wave function of the whole system is given by N independent wave functions, one for each particle. In quantum mechanics, such a state is said to be "separable". In general, however, $\psi(\mathbf{q}_1, \ldots, \mathbf{q}_N)$ is *not* a product, in which case we speak of quantum *entanglement* or we say the "state is entangled".

The main difficulty with thinking in terms of configuration space is that it has very high dimensions, e.g., for a gas of particles, the dimension is of the order of Avogadro's number ($\sim 10^{24}$). It is therefore hard to picture.[10] On the other hand, we really *must* think in terms of configuration space in order to understand quantum mechanics! In later chapters we shall return often to this point, but for the time being we wish to emphasize the meaning of configuration space and to put forward some ideas that can help us to reason in configurational terms.

Our world of experience is spatially located in a three-dimensional space,[11] i.e., bodies like measurement instruments and their pointers occupy regions of space. But the bodies themselves consist of atoms (which themselves consist of smaller objects, but that doesn't matter now) and it is their spatial arrangement, namely their *configuration* that gives the body its shape. This means that the *region in configuration space* determines the body, its shape, its position, and its orientation in space. What must be understood is that different spatial positions of a body are described by disjoint regions in configuration space. If we are concerned with macroscopic bodies (where the number of atoms is of the order of Avogadro's constant), for instance the pointer of a piece of measuring apparatus, then macroscopically different situations (pointer points to the left or to the right) are given by macroscopically disjoint subsets of configuration space (see Fig. 1.1).

[10] The rather beautiful German word "unanschaulich" is often used to describe configuration space, but it has also been abused in the context of quantum mechanics to suggest that one cannot develop any coherent picture of what is going on in the microcosm.

[11] It is often stated in quantum theory that it is of great philosophical importance that in Hilbert space, the space of wave functions, one can select an arbitrary basis in which to express the wave function in coordinates. So there is a position basis, a momentum basis, an energy basis, and whatever we want, and none of those bases is preferred. This is true mathematically. But every human being, even someone working in the field of quantum physics, lives and dies in position space, at least physically. So it is natural enough to find wave functions in the position representation particularly informative.

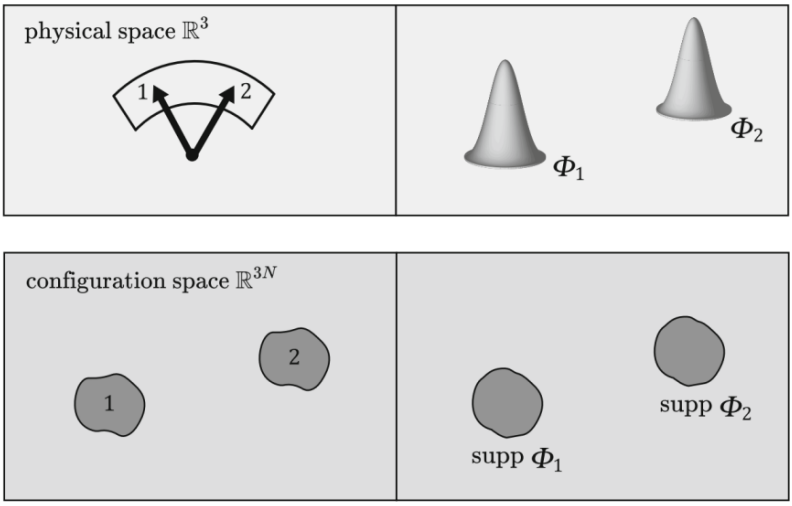

Fig. 1.1 Different macroscopic objects, here two physical pointers consisting of N atoms, where N has the order of magnitude of Avogadro's number, are represented by disjoint regions in the configuration space $\mathbb{R}^{3N}$ (*left*). The wave function of each pointer is concentrated in the corresponding area because of Born's interpretation of the wave function (*right*)

Those configurations that describe a pointer pointing to the left lie in a completely different region of configuration space than those that describe a pointer pointing to the right. And the two regions do not only differ by a few coordinates (or degrees of freedom), but by a huge number of them (order of magnitude 10^{24})—hence the term "macroscopically disjoint".

The wave functions of (macroscopic) bodies are functions on configuration space, and according to Born's rule—because we can see pointers or other macroscopic objects quite sharply—their wave functions are essentially only concentrated in the regions of configuration space which define the shapes, locations, and orientations of the bodies. Above we introduced the notion of support for the set of points where a function is *not* equal to zero. Thus, the wave functions that describe different pointer positions[12] will essentially have macroscopically disjoint supports (see Fig. 1.1). The configuration space of a single particle is the physical space $\mathbb{R}^3$. That is sufficient to understand the double-slit experiment, as discussed. But often in this context the question arises: If we do not close the slits but attach a measuring device designed to determine which slit the particle goes through, what happens

[12]Mathematically trained students will soon realise that if an initial wave function has compact support the time evolved wave function (solving Schrödinger's equation) will immediately have support equal to the whole of configuration space. But the function will be almost zero in most of configuration space. Hence all the following statements about disjoint supports and the associated picture can be taken with a grain of salt, i.e., they should be taken in the sense of physics and not pedantic mathematics.

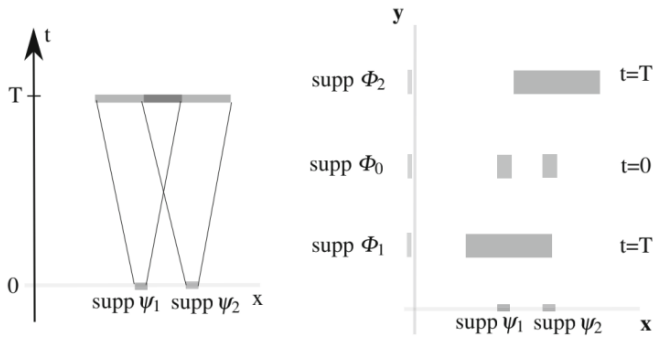

Fig. 1.2 *Left*: Superposition of two initially spatially separated wave functions of a system ($\mathbf{x}$) which spread in time and whose supports overlap at time T. In the overlap region interference occurs. *Right*: In the configuration space of the apparatus ($\mathbf{y}$) and the system, the coupling of a piece of apparatus with pointer wave functions Φ_0 (pointer null setting) and $\Phi_{1,2}$ (displays). The macroscopic wave functions have disjoint supports, which remain disjoint throughout the time evolution. The spreading of the system wave function parts does not lead to any overlap of the supports. Interference is not possible

then? The answer is that *decoherence* takes place and the interference pattern disappears. Simply put, the wave function of the particle becomes entangled with the wave function of the measurement device, and this entangled, high-dimensional wave function no longer interferes. Related to this is the phenomenon that the interference pattern becomes weaker or even disappears completely when very large molecules (say 10^{10} particles, sometimes referred to as Schrödinger cat states) are sent through the double slit.

The disappearance of interference can be explained by Fig. 1.2, which we will now discuss only briefly, because we will repeatedly take up the same reasoning in detail in the following chapters. A measuring device (symbolized by a pointer) is a macroscopic system and a measurement experiment involves an interaction between the system to be measured (e.g., the particle in the double-slit experiment) and the apparatus. The wave function of the total system (particle plus apparatus) evolves according to the Schrödinger equation (1.2) for the total system.

In the following, the reader will have to make some effort to make the connection between the abstract representation and the actual double-slit experiment. We consider a system (coordinates $\mathbf{x}$, dimension m), whose wave function consists of a superposition of two wave functions $\psi_1(\mathbf{x})$ and $\psi_2(\mathbf{x})$ (think of the wave components, which start in each slit as Huygens' spherical waves), coupled to a piece of apparatus (coordinates $\mathbf{y}$, dimension n), represented by the pointer null state $\Phi_0(\mathbf{y})$ and the two possible pointer positions $\Phi_1(\mathbf{y})$ and $\Phi_2(\mathbf{y})$. The coupling is expressed by an interaction potential occurring in the Schrödinger equation and it is

assumed that the corresponding Schrödinger evolution[13] of the total wave function yields

$$\psi_i \Phi_0 \overset{t \longrightarrow T}{\longrightarrow} \psi_i \Phi_i \,, \qquad i = 1, 2 \,, \tag{1.19}$$

where T represents the duration of the measurement. This means that the pointer positions correlate with the wave functions ψ_1 and ψ_2, respectively, the pointers "indicate either ψ_1 or ψ_2", e.g., whether the particle was registered at the upper or at the lower slit. If now the system has initially the wave function $\psi = \psi_1 + \psi_2$ (we ignore normalization constants), then due to linearity (see Remark 1.1) and using (1.19), we obtain

$$\psi \Phi_0 \overset{t \longrightarrow T}{\longrightarrow} \psi_1 \Phi_1 + \psi_2 \Phi_2 \,. \tag{1.20}$$

Look at the corresponding Fig. 1.2 of the supports of the wave functions in configuration space! Actually, this contains everything we need to understand the phenomenon, but to really grasp that, we do need to think deeply about what the figure shows, because thinking in terms of configuration space is not easy. In any case, the conclusion is that the (apparatus + system) wave function parts have disjoint supports after the measurement and therefore pass each other in configuration space without overlapping. Therefore the interference terms (1.18) disappear and we have what is called a decoherent superposition. This does not explain why we always see only one of the two pointer positions (this is the measurement problem), but it does explain the disappearance of the interference pattern on the screen.[14]

Why does something similar happen, namely the suppression of interference, when large molecules are used in the double-slit experiment? Because a large molecule interacts more easily with its surroundings (air, electromagnetic waves, etc.) than a photon or an electron does. Loosely speaking, the environment "measures" the position of the molecule more effectively (like a piece of apparatus) than it does for a photon, for example. That means that interaction is ubiquitous, sometimes very invasive, sometimes less invasive. For macroscopic objects like pointers or cats, the interaction with the environment is very invasive, so interference of superpositions of macroscopic wave functions is practically impossible, because

[13] Starting with a "product wave function" $\Psi(\mathbf{q}) = \Psi(\mathbf{x}, \mathbf{y}) = \psi_i(\mathbf{x})\Phi_0(\mathbf{y})$ expresses the idea of the physical independence of system and apparatus before they interact. In the case of no interaction $V(\mathbf{x}, \mathbf{y}) \approx 0$, the Schrödinger equation (1.2) "factorises" into two independent equations, one for each factor of the product wave function. In that situation, the system and apparatus remain "physically independent". However, it is important to note that $V(\mathbf{x}, \mathbf{y}) \approx 0$ by itself does not imply physical independence because the wave function need not have product structure.

[14] Note that a superposition of wave functions does not necessarily mean that interference takes place. By the same token, if a superposition does not lead to interference, i.e., decoherence takes place, that does not mean that there is no superposition.

the wave parts entangle continuously with the wave function of the environment. Thus the effect of decoherence grows. We may now understand why we do not find such quantum interference effects in our macroscopic world of experiences, i.e., in classical physics, which naturally arose first as a description of nature. That is, we understand why we never have the experience of our interlocutor melting away like a character in an Edgar Allan Poe story.

Macroscopically separated wave components of a superposition run past each other (in configuration space). And the higher the dimensionality of the configuration space, the harder it is for them to "meet". If an experimenter were to set out to perform a truly macroscopic interference experiment (interference of the wave parts in Schrödinger's cat experiment), her task would be as complex as the following: simultaneously reverse the velocity of every gas molecule in a room. This analogy is apt, because it is about the precise control of roughly 10^{24} degrees of freedom. In the quantum mechanics of large systems, these are the phases in the macroscopic wave, in classical mechanics, the velocities of the particles. It can also be said that decoherence is an *irreversible process* in the thermodynamic sense. Since macroscopic wave function parts can no longer interfere and we always "see" only one of the macroscopic parts, one can introduce the famous collapse of the wave function, i.e., pretend that the branch of the wave function that we no longer see simply disappears, which sounds a bit like "putting one's head in the sand". We'll discuss this in more detail in Chap. 2.

1.6 The Classical Limit

This term so often used and heard doesn't actually make much sense. Limits are taken in mathematics for proving theorems and thus do not belong in the world of physics. What exactly does the "classical limit" mean then? The term is used as a shorthand description for physical situations that are more conveniently described by classical mechanics than by quantum mechanics. A bit more needs to be said because we seem to do immensely well with classical physics and it's usually much more difficult to see quantum mechanical effects. Nevertheless, it is generally accepted that quantum mechanics is more fundamental than classical mechanics, i.e., the latter should be included in the new theory as a kind of "limiting case".

It is sometimes said that the classical limit comes about by letting Planck's quantum of action $\hbar$ goes to zero. For example, in the limit $\hbar \to 0$, the commutator of the position and momentum operators becomes the Poisson bracket of position and momentum in classical mechanics. Why is that statement meaningless? Because Planck's constant is fixed at the value $2\pi\hbar = 6.626\,070\,15 \times 10^{-34}\,\mathrm{J\,s}$ no matter what, so it cannot tend anywhere at all. So what is actually intended? There is indeed a lot to be said about a statement like this. For example, it may mean that in certain physical situations, where the evolution of the system is determined by a physical quantity with the dimension of an action (J s) and where the ratio of $\hbar$ to the value of this action is very small, the evolution of the system can be considered classically. And how small is small? So small that the wavy nature which in general determines

Recalling the quantum flux equation (1.5), we have

$$\frac{\partial}{\partial t}\mathbf{j} = -\frac{\mathrm{i}\hbar}{2m}\frac{\partial}{\partial t}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right) .$$

Observing that ψ^* satisfies the complex-conjugated Schrödinger equation [take the conjugate complex of (1.21)], we obtain

$$\frac{\partial}{\partial t}\mathbf{j} = \frac{1}{2m}\big[(H\psi^*)\nabla\psi - \psi^*\nabla(H\psi) + (H\psi)\nabla\psi^* - \psi\nabla(H\psi^*)\big] .$$

Observing further that, for any ψ, φ and using partial integration and the self-adjointness of H,

$$\int \psi^* H\varphi \,\mathrm{d}^3x = \int (H\psi^*)\varphi \,\mathrm{d}^3x ,$$

we get

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle\mathbf{X}\rangle(t) &= \frac{1}{2m}\int \big[\psi^* H\nabla\psi - \psi^*\nabla(H\psi) + \psi H\nabla\psi^* - \psi\nabla(H\psi^*)\big]\mathrm{d}^3x \\
&\overset{(1.21)}{=} \frac{1}{2m}\int \big[\psi^* V\nabla\psi - \psi^*\nabla(V\psi) + \psi V\nabla\psi^* - \psi\nabla(V\psi^*)\big]\mathrm{d}^3x \\
&= \frac{1}{2m}\int \big[-(\nabla V)\psi^*\psi - (\nabla V)\psi\psi^*\big]\,\mathrm{d}^3x \\
&= \frac{1}{m}\big\langle -\nabla V(\mathbf{X})\big\rangle(t) .
\end{aligned}$$

Amazing, isn't it? We get the Newtonian equations "on average". This is a version of the famous Ehrenfest theorem:

$$m\langle\ddot{\mathbf{X}}\rangle(t) = \big\langle -\nabla V(\mathbf{X})\big\rangle(t) . \tag{1.24}$$

We would get the classical limit and with it the identification of the parameter m as the Newtonian mass if we also had

$$\big\langle -\nabla V(\mathbf{X})\big\rangle(t) \approx -\nabla V\big(\langle\mathbf{X}\rangle(t)\big) , \tag{1.25}$$

because then we would get the Newtonian equation of motion for $\langle\mathbf{X}\rangle$. This is where $\psi(\mathbf{x}, t)$ being "concentrated enough" comes into play: recalling the definition (1.22) of the expected value, we see that the fluctuation around the average value is small, so the expected value of the function of the random variable is given approximately by the function of the expected value.

As a final remark, or better, as a warning, we said earlier that the spreading of the wave function is countered by the interaction with the environment, which ensures that the wave packet remains well localized.[16] But why, we may wonder, is the Schrödinger equation still applicable, as we have assumed in the above argument? In fact in such situations, where it is the environment which produces the wave packet, it is no longer valid for the wave packet because it is not a closed system. This means that we still need to justify its approximate validity if we are to ensure that the above computation is well founded. We shall not do that here. We refer instead to *Bohmian Mechanics. The Physics and Mathematics of Quantum Theory*.[17]

1.7 Spin and the Stern–Gerlach Experiment

Another innovation of quantum mechanics is "spin". It is tempting, given the name, to think of a continuously rotating object, but that is not appropriate. Nevertheless, it is easy to say what spin is mathematically. What is not as easy is to explain why, for example, an electron "must have" spin. The "nonclassical bivalence" [as named by Wolfgang Pauli (1900–1958)] of the spin $-1/2$ electron is something really new, something fundamentally quantum mechanical. In the Stern–Gerlach experiment, a silver atom[18] passing through a Stern–Gerlach magnet (which produces a strongly inhomogeneous magnetic field) is deflected like a magnetic dipole which adjusts itself to lie parallel (spin $+1/2$) or antiparallel (spin $-1/2$) to the field gradient (see Fig. 10.1 for an experiment with the Stern–Gerlach setup). A Stern–Gerlach system which is set up "in the z-direction" prepares "spin $+1/2$ particles" or "spin $-1/2$ particles" in the z-direction. Now send a z-spin $+1/2$ particle through a Stern–Gerlach apparatus which is oriented both (1) orthogonal to the z-direction and (2) orthogonal to the flight direction of the particle. Let us say that it is oriented in the y-direction. Then with probability $1/2$ the particle will end up as a y-spin $+1/2$ or a y-spin $-1/2$ particle. If we choose any other direction, only the probabilities for the splits are influenced, but the bivalence remains. This is an experimental fact.

[16] A warning within a warning: the interaction with the environment actually produces a measurement-like situation, as discussed in Chap. 2, i.e., we end up with a superposition of wave packets which are more or less spatially separated. So Chap. 2 is relevant for this too.

[17] D. Dürr and S. Teufel, *Bohmian Mechanics. The Physics and Mathematics of Quantum Theory*. Springer, 2009.

[18] Actually one would like to do the experiment with electrons, but silver atoms have the advantage of being electrically neutral while possessing a net magnetic moment, due to a single outer electron. Therefore, it is as good as an electron as far as spin is concerned. The fact that it is electrically neutral is also good, since the Lorentz force on charged particles does not then contribute.

The wave function of the particle must therefore observe this splitting and the simplest way to achieve this is to assign two degrees of freedom to the ψ-function itself. ψ thus becomes a *spinor wave function*:

$$\psi : \mathbb{R}^3 \to \mathbb{C}^2\,, \quad \psi(\mathbf{x}) = \begin{pmatrix} \psi_1(\mathbf{x}) \\ \psi_2(\mathbf{x}) \end{pmatrix}\,, \tag{1.26}$$

with normalisation

$$\int (\psi, \psi)\,(\mathbf{x})\,\mathrm{d}^3x = \int \begin{pmatrix} \psi_1^*(\mathbf{x}) \\ \psi_2^*(\mathbf{x}) \end{pmatrix} \cdot \begin{pmatrix} \psi_1(\mathbf{x}) \\ \psi_2(\mathbf{x}) \end{pmatrix} \mathrm{d}^3x = \int \left(|\psi_1(\mathbf{x})|^2 + |\psi_2(\mathbf{x})|^2\right) \mathrm{d}^3x = 1\,.$$

A spinor need not be two-dimensional like here. Other dimensions are possible, for example four, for the relativistic Dirac equation [Paul Dirac (1902–1984), see Sect. 11.1.2] for an electron.

The spin degrees of freedom couple to the magnetic field, as the Stern–Gerlach experiment suggests. This is how. The Schrödinger equation for the wave function will be replaced by the so-called *Pauli equation*, where the "potential" V becomes a Hermitian matrix, and such a matrix can always be written in the form

$$V(\mathbf{x})E_2 + \mathbf{B}(\mathbf{x}) \cdot \boldsymbol{\sigma}\,, \quad V(\mathbf{x}) \in \mathbb{R}\,, \quad \mathbf{B}(\mathbf{x}) \in \mathbb{R}^3\,, \quad \boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)\,,$$

where we define the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}\,, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{1.27}$$

The Pauli equation reads

$$\mathrm{i}\hbar\partial_t \psi = \left[\frac{1}{2m}\big(-\mathrm{i}\hbar\nabla - e\mathbf{A}\big)^2 + eV\right]\psi - \mu\boldsymbol{\sigma}\cdot\mathbf{B}\psi\,, \tag{1.28}$$

where we use the following notation:

- $\psi(\mathbf{x}) = \begin{pmatrix} \psi_1(\mathbf{x}) \\ \psi_2(\mathbf{x}) \end{pmatrix}$ is the spinor wave function.
- e a parameter to be identified with the charge of the particle.
- μ is a "coupling constant", called the gyromagnetic factor, which acts like the strength of the magnetic moment.
- $\mathbf{A}$ is the vector potential.
- V is the electrostatic potential.
- $\mathbf{B} = \nabla \times \mathbf{A}$ is the magnetic field.

In the normalisation of the spinor wave function [below (1.26)], we have used the *scalar product in spin space*:

$$(\psi, \psi)(\mathbf{x}) := \left(\psi_1^*(\mathbf{x}), \psi_2^*(\mathbf{x})\right) \begin{pmatrix} \psi_1(\mathbf{x}) \\ \psi_2(\mathbf{x}) \end{pmatrix} = \sum_{i=1,2} \psi_i^*(\mathbf{x}) \psi_i(\mathbf{x}) \,. \tag{1.29}$$

This enters Born's statistical distribution for the position of the particle which, for the Pauli equation, becomes $\varrho^\psi = (\psi, \psi)$. To see this we show that $\varrho^\psi = (\psi, \psi)$ satisfies the continuity equation (1.4) with

$$\mathbf{j}^\psi = \frac{\hbar}{2im}\big[\,(\psi, \nabla\psi) - (\nabla\psi, \psi)\,\big] = \frac{\hbar}{m}\mathrm{Im}\;(\psi, \nabla\psi)\;. \tag{1.30}$$

The required computation is very much analogous to the one leading to (1.4), but replacing the real function V by a Hermitian one, which does not affect the argument.

Remark 1.5 (The Pauli Flux) What we just said is only for starters and we need to add the caveat that there is a debate about what should be considered the right flux, or what should be the physical flux for the Pauli equation. The general idea is that it is not given by (1.30). There is nothing wrong with the computation suggested above, but we must note that the continuity equation contains the divergence of $\mathbf{j}^\psi$, with the consequence that the flux vector is not uniquely fixed by the continuity equation. Any divergence-free vector field can be added to the flux vector, for example, the curl of a vector field $\mathbf{a}(\mathbf{x})$, i.e., $\nabla \times \mathbf{a}(\mathbf{x})$, since by simple vector analysis $\nabla \cdot \nabla \times \mathbf{a}(\mathbf{x}) = \nabla \times \nabla \cdot \mathbf{a}(\mathbf{x}) = 0$. Is that a disaster? No, not really. We can just note this and stay cool. If we wish to mess around with the Schrödinger flux then there are the physical spacetime symmetries which must be respected by additional terms, and this renders the Schrödinger flux pretty much unique.

But in the Pauli situation there does exist a natural additional flux term which is well behaved in all respects and which is sometimes referred to as the Gordon term [Walter Gordon (1883–1939), known from the Klein–Gordon equation] or the *spin flux*. For a one-particle Pauli equation, this reads

$$\frac{\hbar}{2m} \nabla \times (\psi, \boldsymbol{\sigma}\psi)\,.$$

When added to the so-called *convective flux* (1.30), the total flux is

$$\mathbf{J}^\psi{}_{\text{Pauli}} = \frac{\hbar}{m}\left[\mathrm{Im}(\psi, \nabla\psi) + \frac{1}{2}\nabla \times (\psi, \boldsymbol{\sigma}\psi)\right]. \tag{1.31}$$

There are many arguments in the physics literature to justify taking this as the correct Pauli flux. A straightforward and simple one comes with the derivation of the Pauli equation from the relativistic Dirac equation. In particular, the Pauli equation

results from the relativistic Dirac equation in a "non-relativistic limit".[19] The Dirac equation is the relativistic wave equation for an electron, and when Dirac invented it, the Schrödinger wave was replaced by a spinor wave function which naturally had four components, and not two as in the Pauli equation. The Pauli equation arises in this view as a "non-relativistic limit" of the Dirac equation, and when we take the steps required to extract the Pauli equation and consider how this affects the Dirac flux (11.5), which itself satisfies the continuity equation (11.6) for the Dirac equation, the Pauli flux $\mathbf{J}^{\psi}{}_{\text{Pauli}}$ emerges as the "non-relativistic limit" of the Dirac flux.[20]

We mentioned in Sect. 1.4 that the quantum flux provides the statistics of arrival times. In this respect it is interesting to ask whether the arrival times of spin-1/2 particles are determined by the correct Pauli flux. Quantum measurements of first arrival times would provide helpful information, but they are still lacking.[21]

1.7.1 The Pauli Equation and the Stern–Gerlach Experiment

The reasoning behind the appearance of a two-component spinor was based on the experimental fact that the wave function splits in a Stern–Gerlach experiment. This fact will be used in Chap. 10 in the presentation of the famous EPRB argument by Einstein, Podolsky, and Rosen, but in the spin version proposed by Bohm. Therefore we wish to understand at least heuristically how the splitting comes about by examining the Pauli equation.

As noted above, the Stern–Gerlach experiment is done with neutral atoms. Therefore in what follows we shall consider a neutral particle and set $e = 0$ in (1.28). We consider a Stern–Gerlach setup in which the particle crosses the magnetic field in the y-direction, while the magnetic poles are directed in z-direction. We idealize the experiment as taking place in the (y, z)-plane. The magnetic field is then $\mathbf{B} = (0, B_y(y, z), B_z(y, z))$ with $\operatorname{div} \mathbf{B} = 0$. Since the field is inhomogeneous in the z-direction, this implies that there is also an inhomogeneous field in the y-direction, but we shall ignore this effect. Furthermore, we shall see that the deflection which we wish to argue for depends on the derivative $\partial B_z/\partial z$. So to a first approximation, we can look at the first order term in the Taylor expansion of $B_z(y, z)$ around the median value of z (midway between the two poles), which we can set equal to zero. For simplicity, we can also set $B_z(y, 0) = 0$. So we end up with the approximation

$$\mathbf{B} \approx \big(0, B_y(y), b(y)z\big)\,,$$

[19] We need to spell out carefully what is meant by this, but superficially we may consider that in a particular physical situation the speed of the electron is much less than the speed of light.

[20] See, e.g., M. Nowakowski, The quantum mechanical current of the Pauli equation. Am. J. Phys. **67**, 916–919 (1999), and for an overview, W.B. Hodge, S.V. Migirditch, W.C. Kerr, Electron spin and probability current density in quantum mechanics. Am. J. Phys. **82**, 681–690 (2014).

[21] For more on this, see S. Das and D. Dürr, Arrival time distributions of spin-1/2 particles. Sci. Rep. **9**, 2242 (2019), and arXiv:1802.07141.

This notation, where on the left we recognize the projections onto the eigenvectors of σ_z, can be straightforwardly generalized to arbitrary orientations of the Stern–Gerlach magnets.

If the Stern–Gerlach magnet is oriented along $\mathbf{a}$ then the eigenvectors of σ_z will be replaced by the eigenvectors of $\mathbf{a} \cdot \boldsymbol{\sigma}$, usually denoted by $|\uparrow_{\mathbf{a}}\rangle$ and $|\downarrow_{\mathbf{a}}\rangle$. The probabilities of $\mathbf{a}$-spin up and $\mathbf{a}$-spin down will then be computed by projections onto the corresponding spin components, which we shall denote by $|\langle\uparrow_{\mathbf{a}}\,|\psi\rangle|^2$ and $|\langle\downarrow_{\mathbf{a}}\,|\psi\rangle|^2$, respectively.

1.8 Why "Spinors"?

Why is the wave function (1.26) said to be spinor-valued and not simply $\mathbb{C}^2$-valued? Because it is not just two-valuedness which plays a role, but also the *transformation property* under symmetry transformations. The Schrödinger and Pauli wave equations must respect the symmetries of Galilean spacetime. Among these are spatial rotations, and we need to specify how wave functions transform under spatial rotations, so that the equations for them remain invariant. But before we come to that, let us clarify what is actually meant when we speak of the symmetries of Galilean spacetime. For example, speaking about rotations in a pictorial manner, we can say that our physical space has no preferred direction, so that the physical law for the motion of "stuff" must not prefer one spatial direction over another. But how are rotations described mathematically, since mathematics is the language in which the book of nature is written? For that we need to recall some linear algebra.

Three-dimensional physical space is usually represented by the vector space $\mathbb{R}^3$. And when we think of rotations, we naturally think of rotating the direction of a vector through an angle. However, the ingenious discovery of Hermann Grassmann (1809–1877), the father of modern linear algebra, was that the fundamental meaning of a vector must be understood without any reference to angles. The latter are brought in by introducing an extra structure, which one imposes on the vector space. That extra structure is called a *scalar product*. The scalar product gives meaning to angles and rotations. In short, in Grassmann's abstract theory, rotations are linear transformations on the vector space which leave the Euclidean scalar product invariant. These linear transformations are represented by matrices in $\mathbb{R}^{3,3}$, which are orthogonal and special in the sense of having determinant 1. They form a group called $SO(3)$.

Next recall the notion of a vector field $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$, as used in classical physics. At each point $\mathbf{x} \in \mathbb{R}^3$ there sits a vector

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} F_1(\mathbf{x}) \\ F_2(\mathbf{x}) \\ F_3(\mathbf{x}) \end{pmatrix}$$

which carries a geometrical meaning, namely a *direction* and a *length*. This geometrical meaning manifests itself by the transformation property of the vector field. In short, rotating the coordinate system, the coordinate vectors at every point must rotate as well. Consider for example the Newtonian equations of motion of a particle in a force field $\mathbf{F}$:

$$m\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x})\,.$$

If it is to be taken as fundamental, this equation should be invariant under rotations. Hence, if $\mathbf{x}(t)$, $t \in \mathbb{R}$ is a solution of Newton's equation, then $\tilde{\mathbf{x}}(t) := R\mathbf{x}(t)$ must be as well, where $R \in SO(3)$ is a three-dimensional rotation matrix. We can understand this in the sense of a *passive* transformation, in which case $\tilde{\mathbf{x}}(t)$ is physically *the same* solution trajectory, only represented in a coordinate system which has been rotated by R. The Newtonian equation of motion becomes

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\tilde{\mathbf{x}}(t) = mR\ddot{\mathbf{x}} = R\mathbf{F}(\mathbf{x}) = R\mathbf{F}(R^{-1}\tilde{\mathbf{x}}) \stackrel{!}{=} \tilde{\mathbf{F}}(\tilde{\mathbf{x}})\,. \tag{1.38}$$

Hence, if the spatial coordinates are transformed by the rotation matrix R, i.e., $\mathbf{x} \to \tilde{\mathbf{x}} = R\mathbf{x}$, then $\mathbf{F}(\mathbf{x})$ must also be transformed by R at each point, i.e., as $\mathbf{F}(\cdot) \to \tilde{\mathbf{F}}(\cdot) = R\mathbf{F}(R^{-1}\cdot)$. Such a field $\mathbf{F}$ is called a vector field.

The rotation matrices R are representations of the rotation group on the vector space $\mathbb{R}^3$. To make the connection with $\mathbb{C}^2$-valued spinors, we first build a small bridge. We recall that the plane $\mathbb{R}^2$ is isomorphic to the vector space $\mathbb{C}$. A rotation in the plane through an angle α is represented by a special (i.e., determinant 1) orthogonal 2×2 matrix, which means to say that it is in $\in SO(2)$, while in $\mathbb{C}$ the same rotation will be represented by multiplication by $\mathrm{e}^{\mathrm{i}\alpha}$. We thus make the following observation: instead of a two-dimensional representation of the rotations in the plane, we may as well choose a one-dimensional representation in $\mathbb{C}$ given by $\mathrm{e}^{\mathrm{i}\alpha}$, $\alpha \in [0, 2\pi)$. This representation is called the unitary one-dimensional group $U(1)$.

Spinors do something similar, but we need to dig a bit deeper to find a connection between $\mathbb{R}^3$ and $\mathbb{C}^2$. This connection is closely related to the field of quaternions and its isomorphism with $\mathbb{R}^4$, but it is not absolutely necessary to know about that if we use a bit of imagination. It is easy to see that vectors in $\mathbb{R}^3$ can be represented in the following way by Hermitian 2×2 matrices: $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$ can be represented by

$$\mathbf{x} \cong x\sigma_1 + y\sigma_2 + z\sigma_3 = \begin{pmatrix} z & x - \mathrm{i}y \\ x + \mathrm{i}y & -z \end{pmatrix},$$